A new vision for quality assurance and quality control in ecological studies

Sarah E. McCord[x, 1, 4], Sarah H. Burnett[x, 2], Erica M. Christenson[1,3], Ericha M. Courtright[1], Christine M. Laney[5], Connie Maxwell[1], Jason W. Karl[6], Nelson G. Stauffer[+, 1], Shawn W. Salley[1], Amy Slaughter,[1] Justin W. Van Zee[o,1], Nicholas P. Webb[1, 3]

[x]She/Her
[+]He/Him
[o]Other

Quality Assurance

Anticipatory process to prevent errors from being inserted into your data or analysis

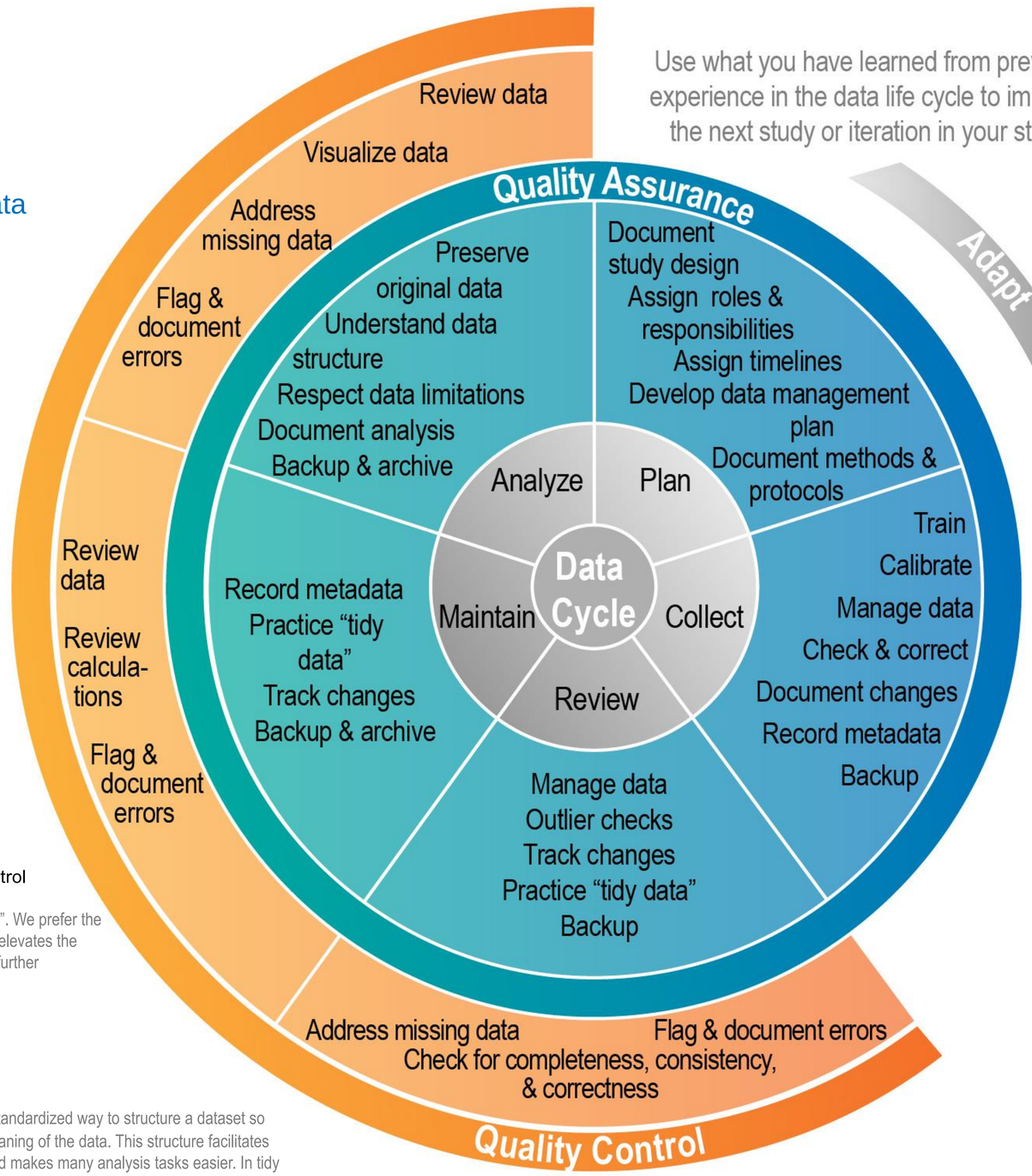
→ Fewer errors in your data or analysis product

Quality Control

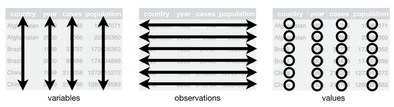
Reactive process to detect, document, and, if possible, fix errors in your data or analysis

→ Validation of quality within data or analysis product

Traditionally, quality assurance and quality control are collectively referred to as "QA/QC". We prefer the annotation "QA & QC" to emphasize the two as distinct and continuous processes. This elevates the preventative steps in QA that may have the most impact on data quality. "QA&QC" also further emphasizes that the two concepts are not necessarily mutually exclusive.



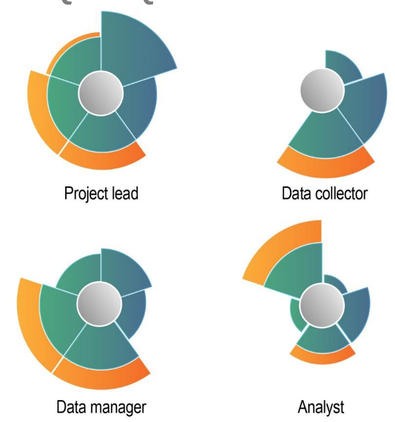
Tidy Data



Wickham and Grolemund. *R for Data Science.* 2017

The tidy data format provides a standardized way to structure a dataset so that it is most aligned with the meaning of the data. This structure facilitates communication about the data and makes many analysis tasks easier. In tidy data, each column represents a variable, each row represents an observation, and each cell represents a value. (Wickham 2014).

Roles in QA & QC



The greatest opportunity to prevent and detect errors may depend on the stage of the data life cycle as well as the role of the individual interacting with the data. The above example of QA and QC effort allocation by role within the Bureau of Land Management's Assessment, Inventory, and Monitoring terrestrial program shows the amount of time an individual with a given role might spend on that step. Cumulatively all roles equal a complete QA&QC data life cycle. One individual may be responsible for more than one role.

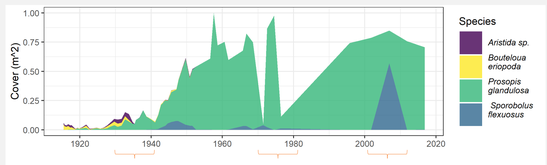
Jornada Quadrat Study 1915-2016

Consistent application of QA & QC is especially critical for long term ecological research. The Jornada Quadrat study is a long term vegetation study, 1915 to present, established to investigate livestock grazing effects on plant community dynamics as well as responses to variable climatic conditions over time. Vegetation monitoring is carried out by charting the basal areas of perennial grasses and canopy areas of shrubs on 122 quadrats. As data collectors and technology shift throughout the study, repeated field work, digitizing the historical data sheets, and analyzing long-term trends all present examples of QA & QC success and challenges.

1935

2016

Stage	Success	Challenge
Plan	✓ Thoughtful sample design along a gradient	▪ Changing technology
Collect	✓ Thorough datasheets ✓ Good notes	▪ Plant name/codes are inconsistent
Review	✓ Manual error checking	▪ Historical data entry process conducive to errors
Maintain	✓ Data are preserved	▪ Multiple data formats ▪ Mildewed paper data sheets
Analyze	✓ Analysis error checking	▪ Different methods of calculating area ▪ Inconsistent vegetation trends



[1]USDA–ARS Jornada Experimental Range, Las Cruces, NM
[2]USDI Bureau of Land Management National Operation Center, Denver, CO
[3]New Mexico State University, Las Cruces, NM
[4]University of Texas-El Paso, El Paso, TX
[5]National Ecological Observatory Network, Boulder, CO
[6]University of Idaho, Moscow, ID

@mccord_se
Sarah.McCord@usda.gov